AMENDMENTS TO THE SPECIFICATION

Please amend page 6, line 14 – page 7, line 20 of the Specification as follows:

The means of identifying homologous sequences and their percentages of homology are well known to those skilled in the art, and include, in particular, the BLAST program, which can be used from the website ~~http://www.ncbi.nlm.nih.gov/BLAST/~~ www.ncbi.nlm.nih.gov/BLAST with the default parameters indicated there. The sequences obtained can then be exploited (e.g. aligned) using, for example, the CLUSTALW program ~~(http://www.ebi.ac.uk/clustalw/)~~ (www.ebi.ac.uk/clustalw) or the MULTALIN program ~~(http://orodes.toulouse.inra.fr/multalin/cgi-bin/multalin.pl)~~ (www.orodes.toulouse.inra.fr/multalin/cgi-bin/multalin.pl),with the default parameters indicated on their websites

Alternatively, the CD-Search program ~~(http://www.ncbi.nih.gov/Structure/cdd/wrpsb.cgi)~~ (www.ncbi.nih.gov/Structure/cdd/wrpsb.cgi) can be used to identify the conserved domains in the protein sequences of *E. coli* or *S. cerevisiae,* and to seek the sequences of other microorganisms presenting the same domain or domains. The conserved domains are recorded in the CDD database (Conserved domain database; Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH. CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Research 31:383-387 (2003)) which groups data of the PFAM or COG type.

The PFAMs (Protein FAMilies database of alignments and hidden Markov models; ~~http://www.sanger.ac.uk/Software/Pfam/)~~ www.sanger.ac.uk/Software/Pfam) are a large collection of alignments of protein sequences. Each PFAM makes it possible to visualize multiple alignments, see protein domains, evaluate distribution among organisms, access other databases, and visualize known protein structures.

The COGs (Clusters of Orthologous Groups of proteins;

~~http://www.ncbi.nlm.nih.gov/COG/)~~ www.ncbi.nlm.nih.gov/COG/) are obtained by comparing protein sequences from 43 fully sequenced genomes representing 30 major phylogenetic lines. Each COG is defined from at least three lines, thus making it possible to identify ancient conserved domains.